Settle for a partial characterization

difficult to obtain or estimate

A full joint distribution function of an arbitrary stochastic process is

A single realization is called a time series

number of possible realizations

A stochastic process is not a single function of time but an infinite

Statistical process describes the time evolution of statistical phenomena

Stationary Processes and Models

ELEC-636: Statistical Signal Processing
\[(\gamma - u)^* \mathbf{r}(u)^T - (\gamma - u, u) \mathbf{r} = \{ \ast [(\gamma - u)^T - (\gamma - u)x][(u)^T - (u)x] \} \mathbf{F} = (\gamma - u, u)\mathbf{c} \]

and

\[\{(\gamma - u)^* x(u)x\} \mathbf{F} = (\gamma - u, u)\mathbf{r}\]

The auto-correlation and auto-covariance are given by

The mean is the mean

\[\{(u)x\} \mathbf{F} = (u)^T\]

Then which may be complex.

\[(W - u)x', \cdots, (I - u)x', (u)x\]

Consider a discrete-time stochastic process
\[
\begin{bmatrix}
(0, \ldots, (Z + IW - ) \ldots, (I + IW - ) \ldots)
\vdots
\vdots
(Z - IW, \ldots, (0, \ldots, (I - ) \ldots, (0, \ldots

\{(u)_H x(u)_{x}\}_{E} = R
\]

A discrete-time stochastic process is wide-sense stationary if

\[\forall \eta \in \mathbb{Z}, \forall u \in I \times IW, \{\eta, \ldots, (Z - u, \ldots, (0, \ldots, (I - u, \ldots, (0, \ldots}
\]

\[\forall \eta \in \mathbb{Z}, \forall u \in I \times IW, (\eta, \ldots, (Z - u, \ldots, (0, \ldots, (I - u, \ldots, (0, \ldots}
\]

\[\forall \eta = (u)_{H}, \forall u \in I \times IW, \eta = (u)_{H}
\]
Properties of the correlation matrices

For a stationary discrete time process

\[ \mathbf{R}^H = \mathbf{R} \text{ (Hermitian)} \]

\[
\begin{bmatrix}
    r^{*}(-1) & r^{*}(0) & \cdots & r^{*}(M) \\
    r^{*}(M) & r^{*}(1) & \cdots & r^{*}(0) \\
    \vdots & \vdots & \ddots & \vdots \\
    r^{*}(2) & \cdots & r^{*}(0) & r^{*}(-1) \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
    r(-1) & r(0) & \cdots & r(M) \\
    r(1) & r(0) & \cdots & r(-1) \\
    \vdots & \vdots & \ddots & \vdots \\
    r(M) & \cdots & r(0) & r(-1) \\
\end{bmatrix}
\]

\[
\left[ \begin{array}{c}
    (\bar{y}_1) \\
    (\bar{y}_2) \\
    \vdots \\
    (\bar{y}_t) \\
\end{array} \right] 
= 
\begin{bmatrix}
    r^{*}(-1) & r^{*}(0) & \cdots & r^{*}(M) \\
    r^{*}(M) & r^{*}(1) & \cdots & r^{*}(0) \\
    \vdots & \vdots & \ddots & \vdots \\
    r^{*}(2) & \cdots & r^{*}(0) & r^{*}(-1) \\
\end{bmatrix} 
\begin{bmatrix}
    \bar{x}_1 \\
    \bar{x}_2 \\
    \vdots \\
    \bar{x}_t \\
\end{bmatrix} 
\]
In this case $R^{-1}$ exists.

$R$ is positive definite if the samples in $x$ are not linearly dependent.

$$0 < \begin{array}{c} a \\ R^a \end{array}$$

and usually

$$0 \preceq \begin{array}{c} a \\ R^a \end{array}$$

For any non-zero vector $a$

$$\begin{bmatrix}
(0) & \cdots & (3 - IW) & (2 - IW) & (1 - IW) \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
(3 - IW) & \cdots & (0) & (1) & (2) \\
(2 - IW) & \cdots & (1) & (0) & (1) \\
(1 - IW) & \cdots & (2) & (1) & (0)
\end{bmatrix} = R$$

The correlation matrix is toplit.
Auto-regressive - moving average - both past input and output used

Moving average - no past model output samples used

Auto-regressive - no past model input samples used

where \( \nu(u) \) is a purely random process.

\[
\begin{array}{c}
(u_x) \\
\text{Linear Filter} \\
\text{Discrete Time} \\
\end{array}
\]

be modeled as

\[
(u_x) (u) \quad \text{physical data observed.}
\]

A model is used to describe the hidden laws governing the generation of

Stochastic Models

We assume that \( x(u) x(u - 1) \ldots \) have statistical dependencies that can
Three possibilities

\[ \text{General Model} \]

\[ \text{ARMA} - \text{mixed AR and MA} \]

\[ \begin{align*}
2. \text{ MA - moving average} \\
1. \text{ AR - auto regressive}
\end{align*} \]

\[ \text{AR part} \]

\[ \text{Linear combination of output of past output} + \text{Model output} \]

\[ \text{MA part} \]

\[ \text{Linear combination of present and past inputs} \]
This is an order \( M \) model and \( \{u\} \) is referred to as the noise term.

\[ (u) = (J - u)x \quad + \cdots + (1 - u)x \quad + \cdots + (u) \]

or

\[ (u) = (J - u)x \quad + \cdots + (1 - u)x \quad + (u) \]

The time series \( \{x\} \) is said to be generated by an AR model if

\[ u = \begin{cases} 0 & \text{otherwise} \\ \varphi & \text{otherwise} \end{cases} \]

\[ \{y\} \sim \mathcal{N}(u, \alpha) \]

\( \alpha \)

For all \( \{u\} \), the input is assumed to be i.i.d. zero mean Gaussian.
\[(z)V = \frac{(z)X}{(z)\Lambda} = (z)\Lambda H\]

Where

If we regard \((u)\alpha\) as the output, then

\[(z)\Lambda = (z)X(z)V\]

Then

\[u_{-z}u_{\alpha}^{0=u_{\infty}} = (z)\Lambda = \{(u)\alpha\}Z\]

\[u_{-z}u_{x}^{0=u_{\infty}} = (z)X = \{(u)x\}Z\]

\[u_{-z}u_{\nu}^{0=u_{\nu}} = (z)V = \{u_{\nu}\}Z\]

Which is a convolution sum. Thus taking Z-transforms

\[(u)\alpha = (\gamma - u)x_{\gamma}^{0=\gamma}\]

Note that we can set \(a = 1\) and write
System is BIBO stable

Impulse response is finite (FIR)

Analyzer is an all zero system

This is called the process analyzer

Legend:
- $z^n$: Sample of input signal
- $x^n$: Sample of output signal
- $\beta$: Weight of noise

Figure 6-36: Statistical Signal Processing
\[
\frac{(z)A}{I} = \frac{(z)A}{(z)X} = (z)^\circ H
\]

If we view \((u)^x\) as the input, then we have the process generator.
\[ W, \ldots, z = u \quad 1 > |ud| \]

Circle, i.e.,

\[
0 = W - zWd + \cdots + z^2d + 1\]

characteristic equation

where \( Wd, \ldots, z^2d \) are the poles of the \( H \) defined as the roots of the

\[
\frac{(1 - zWd - 1) \cdots (1 - z^2d - 1)(1 - z^1d - 1)}{I} = (z)^H
\]

we can factor the denominator and represent \( (z)^H \) in terms of its poles

\[
\frac{u - z^u d}{W} \underbrace{= u}_{I} \quad (z)^V = (z)^H
\]

Since

ELC-636: Statistical Signal Processing
The process generation model is all zero (FIR)

\[ (Y - u)Q^* + \cdots + (I - u)Q + (u) = (u)X \]

The sequence is MA type if

Moving Average Model
The order is

\[(\mathbf{Y} - u)\alpha^{\mathbf{Y}} q + \cdots + (1 - u)\alpha^{\mathbf{q}} + (u)\alpha = (\mathbf{W} - u)\alpha^{\mathbf{W}} p + \cdots + (1 - u)\alpha^{\mathbf{p}} + (u)\alpha\]

past outputs and current/past inputs

In this case, \(\{u(x)\}\) is a mixed process where the output is a function of

Auto-regressive – Moving average model
Note: If $B(z)$ is minimum phase, then it can be represented by an all pole

is perfectly predictable

uncorrelated with $s$

and where $s_t$ is white noise

\[ (y_t - u_t) + q_t = (u_t - x_t) \]

where:

\[ (u_t) + (u_t)x = (u_t)y_t \]

Any stationary discrete time stochastic process $(u_t)$ can be expressed as

**AR** (Autoregressive) system.
\[
\frac{z^{-1}}{1 - \sum_{i=1}^{m} \psi_i z^{-i}} X = (z) \Lambda
\]

In terms of the Z-transform:

\[
(u)(z) \Lambda.
\]

This is a linear constant coefficient difference equation of order \(m\) driven by \(J_N\).

\[
(u)(z) = \sum_{i=1}^{m} \psi_i z^{-i} + (z - u) x z^{-1} + (1 - u) x z^{-1}
\]

Recall that this is generated by \(\{(u)(z)\}\)

Asymptotic statistics of AR processes of AR models are widely used because they are tractable.
The process is asymptotically stationary if \( I > \| u \| \). The homogeneous solution is not stationary. The values depend on the initial conditions:

\[
0 = \nu_{- z} \nu_{\star} \nu_{d} + \cdots + \zeta_{- z} \zeta_{\star} \nu_{d} + \nu_{1} \nu_{d} + I
\]

where \( p_{1}, \ldots, p_{k} \) are the roots of \( \nu_{d}, \ldots, \nu_{d} \).

This has stationary statistics. The homogeneous solution is of the form:

\[
\frac{\nu_{d} \nu_{d} \theta_{d} + \cdots + \zeta_{d} \zeta_{d} \theta_{d} + \nu_{1} \nu_{d} \theta_{d}}{\nu_{d} \nu_{d} \theta_{d} + \cdots + \zeta_{d} \zeta_{d} \theta_{d} + \nu_{1} \nu_{d} \theta_{d}} = (u)^{\circ} x
\]

where \( \nu_{d} \) is taken as the delay operator:

\[
(u)^{\nu(z)} \theta_{H} = (u)^{d} x
\]

The particular solution is the result of driving with \( (z)^{\theta} \theta_{H} \) with \( (u)^{\nu} \theta \). The homogeneous solution:

\[
\{ (u)^{d} x \} + \{ (u)^{\circ} x \} = (u) x
\]

Inverting this leads to the solution.
\[ 0 < \gamma \quad \text{for} \quad 0 = (1 - u) \ast x(u)\alpha \{ \mathcal{E} \}
\]
\[ (\gamma - 1) \alpha = (1 - u) \ast x(\gamma - u) \{ \mathcal{E} \}
\]

Note that

\[ \{ (1 - u) \ast x(u)\alpha \} \mathcal{E} = \left\{ (1 - u) \ast x(\gamma - u) x^{\gamma} \alpha \bigcap_{\mathcal{W}}^{0 = \gamma} \right\} \mathcal{E} \]

\{ \} \mathcal{E} \quad \text{and take} \quad (1 - u) \ast x \quad \text{both sides by} \quad (1 - u) \ast \alpha \]

\[ (u) \alpha = (\gamma - u) x^{\gamma} \alpha \bigcap_{\mathcal{W}}^{0 = \gamma} \]

Recall that an AR process can be written as

\underline{Correlation of a Stationary AR Process}
Thus $\lim \limits_{t \to \infty} (\mu_1, \ldots, \mu_p) > |w_0|$, which are identical to the roots of the AR characteristic equation:

$$0 = \sum_{l=1}^{\infty} \alpha_l m - \cdots - \sum_{l=1}^{\infty} \beta_l z^l m - \sum_{l=1}^{\infty} \gamma_l \varphi_l m$$

where $p$ is the $l$th root of

$$\sum_{l=1}^{\infty} \alpha_l m^l = (\mu_1, \ldots, \mu_p)$$

Note that this also has the solution

$$\sum_{l=1}^{\infty} \alpha_l m^l = \sum_{l=1}^{\infty} \beta_l z^l m + (1 - \gamma) \varphi_l m = (\gamma, \ldots, \gamma)$$

Thus the auto-correlation of the AR process satisfies

$$\sum_{l=1}^{\infty} \alpha_l m^l = (\gamma, \ldots, \gamma)$$

Thus, the above becomes

$$0 < l \quad \text{for} \quad 0 = (\gamma - \gamma) \sum_{l=1}^{\infty} \alpha_l m^l$$
\[(\mathbf{I} - \mathbf{W})_\ast \mathbf{u}_\ast \mathbf{m} + \cdots + (\mathbf{I})_\ast \mathbf{u}_\ast \mathbf{m} + (0)_\ast \mathbf{u}_\ast \mathbf{m} = (\mathbf{I})_\ast \mathbf{u}\]

or

\[(\mathbf{W} - \mathbf{I})_\ast \mathbf{u}_\ast \mathbf{m} + \cdots + (\mathbf{I} - \mathbf{I})_\ast \mathbf{u}_\ast \mathbf{m} + (0)_\ast \mathbf{u}_\ast \mathbf{m} = (\mathbf{I})_\ast \mathbf{u}\]

Lettting \( \mathbf{y} = \mathbf{I} = 1 \) and using the fact \( \mathbf{I} = \mathbf{I} \)

\[(\mathbf{W} - \mathbf{I})_\ast \mathbf{u}_\ast \mathbf{m} + \cdots + (\mathbf{I} - \mathbf{I})_\ast \mathbf{u}_\ast \mathbf{m} + (0)_\ast \mathbf{u}_\ast \mathbf{m} = (\mathbf{I})_\ast \mathbf{u}\]

Recall

\[(\mathbf{W})_\ast \mathbf{u}, \dots, (\mathbf{I})_\ast \mathbf{u}, (0)_\ast \mathbf{u}\]

These parameters can be determined by the auto-correlation values:

- Variance of \( \mathbf{u} \): \( \sigma^2 \)
- AR coefficients: \( a_1, a_2, \dots, a_p \)

An AR model of order \( p \) is completely specified by the Yule-Walker Equations.
\[ L[(\xi - I)w, \cdots, (I)w, (0), (I)w, (2)w, (0)] \] = \\
\[ (\xi - I)w_m + \cdots (I)w_m + (0)w_m + (I)w_m + (2)w_m = (\xi)w \] \\
\[ (I - I)w_m + \cdots (I)w_m + (0)w_m + (I)w_m = (I)w \]

Similarly, for \( \xi = 1 \)

\[ L[(\zeta - I)w, \cdots, (I)w, (0), (I)w] \] = \\
\[ (\zeta - I)w_m + \cdots (I)w_m + (0)w_m + (I)w_m = (\zeta)w \] \\
\[ (I - \zeta)w_m + \cdots (I)w_m + (0)w_m + (I)w_m = (I)w \]

Now let \( \zeta = 2 \)

\[ [w_m, \cdots, z_m, \bar{m}] = L \]

Where

\[ L[(I - I)w, \cdots, (I)w, (0)] \] = \\
\[ (I - I)w_m + \cdots (I)w_m + (0)w_m = (I)w \]

or taking the complex conjugate
\[ W, \ldots, W = y \]

\[ \mathbf{H}^{-1} = W \]

where

\[ \mathbf{H} = W \]

coefficients — assuming \( \mathbf{H} \) is nonsingular

This gives the auto-correlation values, we can uniquely determine the AR

\[ \mathbf{H} \]

or more compactly

\[
\begin{bmatrix}
(W)_{*d} \\
(\mathcal{E})_{*d} \\
(\mathcal{Z})_{*d} \\
(\mathcal{I})_{*d}
\end{bmatrix}
= \begin{bmatrix}
\rho_{m} \\
\varepsilon_{m} \\
\varphi_{m} \\
\varpi_{m}
\end{bmatrix}
\begin{bmatrix}
\begin{pmatrix}
0 \ldots (\mathcal{E} - W)_{*d} \ldots (\mathcal{Z} - W)_{*d} \ldots (1 - W)_{*d}
\end{pmatrix} \\
\begin{pmatrix}
(\mathcal{E} - W)_{*d} \ldots (0)_{*d} \ldots (1)_{*d} \ldots (0)_{*d}
\end{pmatrix} \\
\begin{pmatrix}
(\mathcal{Z} - W)_{*d} \ldots (1)_{*d} \ldots (0)_{*d} \ldots (1)_{*d}
\end{pmatrix} \\
\begin{pmatrix}
(1 - W)_{*d} \ldots (0)_{*d} \ldots (1)_{*d} \ldots (0)_{*d}
\end{pmatrix}
\end{bmatrix}
\]

or in matrix form

\[ \begin{bmatrix}
(\mathcal{E})_{*d} \\
(\mathcal{Z})_{*d} \\
(\mathcal{I})_{*d}
\end{bmatrix}
= \begin{bmatrix}
\rho_{m} \\
\varepsilon_{m} \\
\varphi_{m} \\
\varpi_{m}
\end{bmatrix}
\begin{bmatrix}
(\mathcal{E} - W)_{*d} \ldots (0)_{*d} \ldots (1 - W)_{*d}
\end{bmatrix}
\]
\[
(\gamma)_{\mathcal{A}}^N_{\mathcal{B}}^0 = \mathcal{A} \\
\Rightarrow \{ (u)_{\mathcal{B}}^\ast (u)_{\mathcal{A}} \} \mathcal{A} = (\gamma - u)_{\mathcal{B}}^\ast \mathcal{B}^N_{\mathcal{B}}^0
\]

Thus

\[
\mathcal{I} \mathcal{W} = \gamma, 0 = \{ (\gamma - u)_{\mathcal{B}}^\ast \mathcal{B}^N_{\mathcal{B}}^0 \}
\]

But also

\[
\{[(u)_{\mathcal{B}}^\ast + (\mathcal{I} - u)\mathcal{W}^N_{\mathcal{B}}^\ast + \mathcal{Z} - u)\mathcal{Z}^N_{\mathcal{B}}^\ast + (I - u)\mathcal{I}^N_{\mathcal{B}}^\ast + (\mathcal{I} - u)\mathcal{I}^N_{\mathcal{B}}^\ast ](u)_{\mathcal{A}} \} \mathcal{A} =
\]

\[
\{(u)_{\mathcal{B}}^\ast x(u)_{\mathcal{A}} \} \mathcal{A} = (\gamma - u)_{\mathcal{B}}^\ast \mathcal{B}^N_{\mathcal{B}}^0
\]

Let

\[
0 = \mathcal{I}
\]

\[
\{(1 - u)_{\mathcal{B}}^\ast x(u)_{\mathcal{A}} \} \mathcal{A} = (\gamma - 1)_{\mathcal{B}}^\ast \mathcal{B}^N_{\mathcal{B}}^0
\]

\[
\{(1 - u)_{\mathcal{B}}^\ast x(\gamma - u)_{\mathcal{B}}^\ast \} \mathcal{A} = \{ (1 - u)_{\mathcal{B}}^\ast x(\gamma - u)_{\mathcal{B}}^\ast \} \mathcal{A}
\]

Recall that we started with

\[
\{(u)_{\mathcal{B}}^\ast \}
\]

Lastly, we must determine the variance of the driving sequence \(v(t)\).
\((u) n = (z - u) x \tilde{n} + (1 - u) x \tilde{v} + (u) x\)
\[
I \geq z_0 I - I_0 \\
I_0 - z_0 I \geq I - I_0 \\
I_0 + z_0 I \geq I - I_0 \\
(\overline{z_0 I - I_0}) \frac{z_0}{I} = \bar{z}d, I \leq \bar{z}d \\
0 = z_{-z_0 I + I_{-1} I + I}
\]

has characteristic equation
Recall that the auto-correlation can be expressed as
The characteristics of the AR process vary in a related fashion to the pole placements.
Let $\theta$ be the estimated model (AR/MA/ARMA) order $m$ parameters.

Let $\mathcal{L}[\theta_1^{m_1}, \ldots, \theta_m^{m_m}] = \theta$

Information theoretic criteria: estimated.

In addition to estimating model parameters the model order must also be estimated. A model is typically estimated from a finite set of observation data.

Model Order Solution
\[ \text{AIC} = (w) \mu \left( 2 \log \left( \sum_{N}^{w=1} \mu \theta \mid \frac{1}{2} \right) \right) - 2 \log \left( \sum_{N}^{w=1} \mu \theta \right) \]

Parameter cost function

Parameter cost function

Always decreasing

Let \( \text{AIC} \) model order be given by \( m \) that minimizes

Then the AIC model order is given by \( m \) that minimizes parameters.

To be the logarithm of the maximum likelihood estimates of the model

\[ \left( \mu \theta \mid \frac{1}{2} \right) \]

Set

\[ \text{defined by} \]

\[ \text{Let} \]

\[ \mu \theta \]

\( \mu \theta \mid \frac{1}{2} \)