4. MODELS FOR STATIONARY TIME SERIES

Model = Stochastic discrete recursive (or ARMA) linear system:

4.1. Motivation

$$\mathcal{U}=\{U_t\} \longrightarrow \boxed{\quad LS_h \quad} \longrightarrow X=\{X_t\} \qquad t\in\mathbb{Z}$$

$U_t, X_t \in L_2(\Omega, \mathcal{A}, P) =: L_2$ ... r.v. with finite second moments

$\mathcal{U}, X : \mathbb{Z} \to L_2$ ... input, output = time series with finite second moments.

$h \in \ell_1$ (stability) ... impulse response

L_2 space

$L_2 := \{Y \mid E|Y|^2 < \infty\}$ ... inner-product linear space $\underbrace{}_{}$ over $\mathbb{R}$ (or $\mathbb{C}$) which is complete (Hilbert space)

$\mathbb{R} \subset L_2$

① $X, Y \in L_2$: $\langle X,Y\rangle := EXY$ ... inner-product

② $X \in L_2$: $\|X\| = \sqrt{\langle X,X\rangle} = \sqrt{E|X|^2}$ ... norm

Schwarz inequality: $|\langle X,Y\rangle| \le \|X\|\cdot\|Y\|$

③ $X, X_n \in L_2$: $X_n \xrightarrow{2} X \overset{\text{def.}}{\equiv} \|X - X_n\| \to 0 \equiv E|X-X_n|^2 \to 0$ for $n\to\infty$

... mean-square convergence

$$X \overset{2}{=} \sum_{\substack{n=-\infty \\ (n=0)}}^{\infty} X_n \overset{\text{def.}}{\equiv} \sum_{\substack{n=-N \\ (n=0)}}^{N} X_n \xrightarrow{2} X \text{ for } N\to\infty$$

... mean-square summation

④ $X, Y \in L_2$: $EX = \langle X, 1\rangle$, $\operatorname{cov}(X,Y) = E(X-EX)\cdot(Y-EY) = \langle X-EX, Y-EY\rangle$

$\operatorname{var} X = \sigma_X^2 = E|X-EX|^2 = \|X-EX\|^2$

⑤ Continuity: $X_n \xrightarrow{2} X, Y_n \xrightarrow{2} Y \Rightarrow \langle X_n, Y_n\rangle \to \langle X,Y\rangle \Rightarrow \|X_n\| \to \|X\|$

Hence by ④: $EY_n \to EX$, $\operatorname{cov}(X_n, Y_n) \to \operatorname{cov}(X,Y)$

$\operatorname{var} Y_n \to \operatorname{var} X$

$$\boxed{X = LS_h(\mathcal{U})} \equiv \boxed{X_t \overset{2}{=} \sum_{\substack{k=-\infty \\ (k=0)}}^{\infty} h_k U_{t-k}} \equiv \boxed{X_t = H(B)U_t} \text{ by 3.20}$$

Goals to show

(1) $\mathcal{U}$ stationary $\Rightarrow$ X stationary

(2) X arbitrary stationary, $\mu_X = 0$ $\Rightarrow$ $\exists$ causal and stable LS_h, $h\in\ell_1$ and $\{Z_t\} \sim WN(0,\sigma^2)$ such that $X_t = LS_h(Z_t)$.

(3) LS_h may be approximated with arbitrary precision (matching of acf) by a recursive linear system (ARMA-LS).

4.2. Stability of stochastic discrete LS: $X = LS_h(U)$

$\|U_t\| \le C_1 < \infty \ \forall t \Rightarrow \exists C_2 < \infty : \|X_t\| \le C_2 \ \forall t$

(Bounded inputs produce bounded outputs)

Rephrasing boundedness in stochastic setting:

4.2.1. Lemma

$\|X_t\| \le C < \infty \ \forall t$ iff $\exists \mu_X, \sigma_X < \infty : |\mu_X(t)| \le \mu_X \ \& \ |\sigma_X(t)| \le \sigma_X$.

($\{X_t\}$ is bounded iff $\{X_t\}$ has bounded both mean and variance)

Proof:

$\Rightarrow$: $\|X_t\| \le C \Rightarrow |\mu_X(t)| = |E X_t| = |\langle X_t, 1\rangle| \le \|X_t\| \cdot \underbrace{\|1\|}_{1} \le C =: \mu_X$

$\Rightarrow |\sigma_X(t)|^2 = E|X_t|^2 - |EX_t|^2 \le E|X_t|^2 = \|X_t\|^2 \le C^2 =: \sigma_X^2$

$\Leftarrow$: $\|X_t\|^2 = E|X_t|^2 = |\sigma_X(t)|^2 + \underbrace{|EX_t|^2}_{|\mu_X(t)|^2} \le \sigma_X^2 + \mu_X^2 =: C^2$.

4.2.2. Corollary (Rephrasing stability in stochastic setting)

LS_h is stable iff $(X = LS_h(U), |\mu_U(t)| \le \mu_U < \infty, |\sigma_U(t)| \le \sigma_U < \infty \ \forall t \Rightarrow$
$\Rightarrow \exists \mu_X, \sigma_X < \infty : |\mu_X(t)| \le \mu_X \ \& \ |\sigma_X(t)| \le \sigma_X \ \forall t)$

4.3. Main Convergence Theorem

If $h \in l_1$ (i.e. $\sum_k |h_k| < \infty$) and $\sup_t E|U_t|^2 < \infty$, then $\sup_t E|U_t| < \infty$ (i.e. $\{U_t\}$ is bounded), $X_t :\overset{2}{=} \sum_k h_k U_{t-k}$ converges and defines a stable stochastic linear system.

Proof

I. $C^2 := \sup_t E|U_t|^2 = \sup_t \||U_t|\|^2 = \sup_t \|U_t\|^2 \Rightarrow \|U_t\|^2 \le C^2 \overset{4.2.1}{\Rightarrow} E|U_t| = |EU_t| =$
$= |\mu_U(t)| \le \mu_U < \infty \Rightarrow \sup_t E|U_t| \le \mu_U < \infty$.

II. $S_n := \sum_{\substack{j=-n \\ (j=0)}}^{n} h_j U_{t-j}$, $n > m$ arbitrary $\Rightarrow \|S_n - S_m\| = \|\sum_{m<|j|\le n} h_j U_{t-j}\| \le$

$\le \sum_{m<|j|\le n} |h_j| \underbrace{\|U_{t-j}\|}_{\le C} \le C \cdot \sum_{m<|j|\le n} |h_j| \to 0$ for $n, m \to \infty$ because

$\sum_j |h_j| = \lim_{n\to\infty} \sum_{\substack{j=-n \\ (j=0)}}^{n} |h_j| \Rightarrow \{\sum_{\substack{j=-n \\ (j=0)}}^{n} |h_j|\}$ is convergent and therefore a Cauchy sequence. Then $\{S_n\}$ is Cauchy seq. in L_2, which is convergent by completeness of L_2. Its limit is exactly X_t.

III. Stability: $\sum_j |h_j| < \infty$, $\|U_t\| \le C \ \forall t \Rightarrow \forall t$:

$$\|X_t\| = \|\sum_j h_j U_{t-j}\| = \|\lim_{n\to\infty} s_n\| = \lim_{n\to\infty} \|s_n\| =$$

$$= \lim_{n\to\infty} \|\sum_{\substack{j=-n \\ (j=0)}}^{n} h_j U_{t-j}\| \le \lim_{n\to\infty} \sum_{\substack{j=-n \\ (j=0)}}^{n} |h_j| \underbrace{\|U_{t-j}\|}_{\le C} \le C \cdot \sum_j |h_j| < \infty.$$

4.4. Corollary

If $h \in \ell_1$ and $U = \{U_t\}$ is stationary with mean μ_U and a.c.f. $\gamma_U(\cdot)$ then

(1) $X_t \overset{2}{=} \sum_j h_j U_{t-j}$ converges $\forall t \in \mathbb{Z}$.

(2) $\{X_t\}$ is stationary: $\boxed{\mu_X = \mu_U \sum_j h_j \text{ and } \gamma_X(h) = \sum_j \sum_k h_j h_k \gamma_U(h-j+k),\ h \ge 0.}$

Proof

(1) $\{U_t\}$ stationary $\Rightarrow E|U_t|^2 =: C^2$ (constant) $\Rightarrow \sup_t E|U_t|^2 = C^2 < \infty$ and convergence follows by 4.3.

(2) $X_t \overset{2}{=} \sum_j h_j U_{t-j} \Rightarrow \sum_{\substack{j=-n \\ (j=0)}}^{n} h_j U_{t-j} \overset{2}{\to} X_t \Rightarrow E\left(\sum_{\substack{j=-n \\ (j=0)}}^{n} h_j U_{t-j}\right) =$

$$= \sum_{\substack{j=-n \\ (j=0)}}^{n} h_j \underbrace{E U_{t-j}}_{\mu_U} = \mu_U \sum_{\substack{j=-n \\ (j=0)}}^{n} h_j \longrightarrow \mu_U \sum_j h_j \underset{\uparrow \text{ by contin. of } E}{=} E X_t$$

$$\left[\text{shortly: } EX_t = E\left(\sum_j h_j U_{t-j}\right) = \sum_j h_j \underbrace{E U_{t-j}}_{\mu_U} = \mu_U \sum_j h_j =: \mu_X\right]$$

Analogically using short form:

$$\gamma_X(h) = \operatorname{cov}(X_{t+h}, X_t) = \operatorname{cov}\left(\sum_j h_j U_{t+h-j}, \sum_k h_k U_{t-k}\right) =$$

$$= \sum_j \sum_k h_j h_k \underbrace{\operatorname{cov}(U_{t+h-j}, U_{t-k})}_{\gamma_U(t+h-j-(t-k))} = \sum_j \sum_k h_j h_k \gamma_U(h-j+k)$$

4.4.1. Corollary

In particular: $\boxed{\sigma_X^2 = \gamma_X(0) = \sum_j \sum_k h_j h_k \gamma_U(k-j)}$ and $\boxed{\sigma_X^2 \le \sigma_U^2 \left(\sum_j |h_j|\right)^2}$

Proof

$$\sigma_X^2 = |\sigma_X^2| \le \sum_j \sum_k |h_j||h_k| \underbrace{|\gamma_U(k-j)|}_{\le \gamma_U(0) = \sigma_U^2} \le \sigma_U^2 \sum_j \sum_k |h_j||h_k| = \sigma_U^2 \left(\sum_j |h_j|\right)^2.$$

4.5. Remark

(1) The time series $\{X_t\}$ of theorem 4.3. is also said to be a time series generated by time series $\{U_t\}$.

(2) Such $\{X_t\}$ is said to be causal if $h_k = 0$ for $k < 0$ (cf. Th. 3.7(2))

(3) The stability condition $h \in \ell_1$ in theorem 4.3. may be replaced by equivalent condition on $\mathcal{H}(z)$ according to Theorem 3.14 (2) or (3).

BEST LINEAR PREDICTION

4.6. Theorem (The Projection Theorem – see [BD93, §2.3])

If $\mathcal{L}$ is a closed linear subspace of a Hilbert space $\mathcal{H}$ (e.g. $\mathcal{H} = L_2$) and $x \in \mathcal{H}$, then

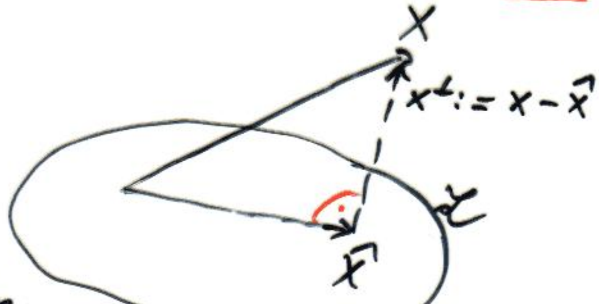
(i) there exists a unique element $\hat{x} \in \mathcal{L}$ such that

$$\|x - \hat{x}\| = \inf_{y \in \mathcal{L}} \|x - y\| \qquad (*)$$

and

(ii) some $\hat{x} \in \mathcal{L}$ has the minimality property $(*)$ iff $x - \hat{x} \perp \mathcal{L}$, i.e. $x - \hat{x} \perp y$ (or $\langle x - \hat{x}, y \rangle = 0$) $\forall y \in \mathcal{L}$. We write $\hat{x} = P_{\mathcal{L}} x$.

4.7. Definition



Let $X_0 := 1, X_1, \dots, X_n \in L_2$ and let $\mathcal{L} := \mathcal{L}(X_0, X_1, \dots, X_n)$ be a linear subspace in L_2 spanned by $X_0, X_1, \dots, X_n$.

Given $X \in L_2$, then $\hat{X} = P_{\mathcal{L}} X$ is called the best (mean square) linear prediction of X in terms of $X_0, X_1, \dots, X_n$.

$\dim \mathcal{L} \leq n+1$ is finite $\Rightarrow$ $\mathcal{L}$ is closed $\Rightarrow$ $P_{\mathcal{L}}$ is well-defined.

As $\hat{X} \in \mathcal{L}(X_0, X_1, \dots, X_n)$ there exists $\underline{\beta}_0 := [\beta_0, \beta_1, \dots, \beta_n]^T \in \mathbb{R}^{n+1}$ such that

$$(4.1a) \qquad \boxed{\hat{X} = \beta_0 \underbrace{X_0}_{1} + \beta_1 X_1 + \dots + \beta_n X_n}$$

Clearly $\underline{\beta}_0$ is unique iff $X_0, X_1, \dots, X_n$ are linearly independent, (i.e. if they constitute a basis of $\mathcal{L}$)

4.8. Theorem

The vector β_0 of (4.1a) is obtained as a solution to the following system of linear equations (so-called normal equations):

$$(4.1b)\quad \boxed{\underbrace{(EX_0X_i)}_{\langle X_0,X_i\rangle}\beta_0+\underbrace{(EX_1X_i)}_{\langle X_1,X_i\rangle}\beta_1+\dots+\underbrace{(EX_nX_i)}_{\langle X_n,X_i\rangle}\beta_n=\underbrace{EXX_i}_{\langle X,X_i\rangle}\quad i=0,1,\dots,n}$$

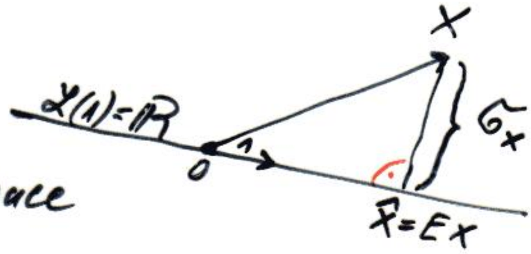
Proof:
By 4.6.(ii): $\hat{X}=P_{\mathcal{L}}X \Leftrightarrow X-\hat{X}\perp y\ \forall y\in\mathcal{L} \Leftrightarrow X-\hat{X}\perp X_i\ \forall i=0,1,\dots,n$

$\Leftrightarrow \underbrace{\langle X-\hat{X},X_i\rangle}_{\langle X,X_i\rangle-\langle\hat{X},X_i\rangle}=0\ \forall i=0,1,\dots,n \Leftrightarrow \langle\hat{X},X_i\rangle=\langle X,X_i\rangle\ \forall i=0,1,\dots,n$

$\Leftrightarrow \langle\sum_{j=0}^{n}\beta_jX_j,X_i\rangle=\langle X,X_i\rangle\ \forall i=0,1,\dots,n \Leftrightarrow \sum_{j=0}^{n}\beta_j\langle X_j,X_i\rangle=\langle X,X_i\rangle\ \forall i=0,1,\dots,n$

which is (4.1b). □

4.9. Example ($n=0$)



$\mathcal{L}:=\mathcal{L}(X_0)=\mathcal{L}(1)=\mathbb{R}$ ... linear subspace of all (real) constants in L_2.

Then (4.1b) is one equation: $\underbrace{(E\underbrace{X_0X_0}_{1})}_{1}.\beta_0=EX.\underbrace{X_0}_{1}\Rightarrow\beta_0=EX\Rightarrow$

$\Rightarrow\hat{X}=\beta_0.\underbrace{X_0}_{1}=EX.$

$\|X-EX\|=\sqrt{E(X-EX)^2}=\sigma_X$ is the minimal distance of X from $\mathbb{R}$.

4.10. Corollary of 4.8.

If $EX=EX_1=\dots=EX_n=0$ then $\hat{X}=P_{\mathcal{L}(X_0,X_1,\dots,X_n)}X=P_{\mathcal{L}(X_1,\dots,X_n)}X$, $\hat{\beta}_0=0$ and (4.1b) attains the form:

$$(4.1c)\quad \boxed{(EX_1X_i)\beta_1+\dots+(EX_nX_i)\beta_n=EX.X_i,\quad i=1,\dots,n}$$

which with $\mathbb{X}:=[X_1,\dots,X_n]^T$ may be rewritten in matrix form:

$\boxed{(\operatorname{var}\mathbb{X})\beta=\operatorname{cov}(X,\mathbb{X})^T}$ where $\beta=[\beta_1,\dots,\beta_n]^T$.

Proof: $i=0$ in (4.1b) $\Rightarrow \underbrace{(E\underbrace{X_0X_0}_{1})}_{1}\beta_0+\underbrace{(EX_1.\underbrace{X_0}_{1})}_{0}\beta_1+\dots+\underbrace{(EX_n\underbrace{X_0}_{1})}_{0}\beta_n=\underbrace{EX\underbrace{X_0}_{1}}_{=0}$

$\Rightarrow\beta_0=0\Rightarrow(EX_0X_i)\beta_0=0$ for $i=1,\dots,n$ yielding (4.1c) which is a system of normal eqs. solving $P_{\mathcal{L}(X_1,\dots,X_n)}X$.

4.11 Corollary (Best linear prediction in stationary time series)

$X = \{X_t \mid t \in \mathbb{Z}\}$ stationary with $\mu_X = 0$ and autocov. function $\gamma(\cdot)$. For given $k \in \mathbb{N}$ let $\hat{X}^{(k)}_{t+k}$ be the best (k-step) linear prediction of X_{t+k} in terms of n previous terms $X_t, X_{t-1}, \dots, X_{t+1-n}$ in the form $\hat{X}^{(k)}_{t+k} = \sum_{j=1}^{n} \phi^{(k)}_{n,j} X_{t+1-j}$. Then the vector $\underline{\phi}^{(k)}_n = [\phi^{(k)}_{n1}, \dots, \phi^{(k)}_{nn}]^T$ is found as a solution to the so-called Yule-Walker system of linear equations:

$$(4.2) \qquad \underbrace{\begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(n-2) \\ \vdots & & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \cdots & \gamma(0) \end{bmatrix}}_{=:\, \Gamma_n} \cdot \underbrace{\begin{bmatrix} \phi^{(k)}_{n1} \\ \phi^{(k)}_{n2} \\ \vdots \\ \phi^{(k)}_{nn} \end{bmatrix}}_{\underline{\phi}^{(k)}_n} = \underbrace{\begin{bmatrix} \gamma(k) \\ \gamma(k+1) \\ \vdots \\ \gamma(k+n-1) \end{bmatrix}}_{\underline{\gamma}^{(k)}_n}$$

Proof: $\mu_X = 0 \Rightarrow EX_{t+k} = EX_t = \dots = EX_{t+1-n} = 0 \overset{4.10}{\Rightarrow}$ $\hat{X}^{(k)}_{t+k} = P_{\mathcal{L}(X_t, X_{t-1}, \dots, X_{t+1-n})} X_{t+k}$ and (4.2) is obtained from (4.1c) with replacements $X \rightsquigarrow X_{t+k}$, $X_j \rightsquigarrow X_{t+1-j}$, $\beta_j \rightsquigarrow \phi^{(k)}_{nj}$:

$\operatorname{var} \mathbb{X} \rightsquigarrow [E X_{t+1-i} X_{t+1-j}] = [\operatorname{cov}(X_{t+1-i}, X_{t+1-j})] = [\gamma(\underbrace{t+1-i-(t+1-j)}_{j-i})] = \Gamma_n$ in view of symmetry $\gamma(-h) = \gamma(h)$.

$\operatorname{cov}(X, \mathbb{X})^T \rightsquigarrow [E X_{t+k} X_{t+1-j}] = [\operatorname{cov}(X_{t+k}, X_{t+1-j})] = [\gamma(\underbrace{t+k-(t+1-j)}_{k+j-1})] = \underline{\gamma}^{(k)}_n$.

4.12. Theorem (Mean square error of the best (k-step) linear prediction)

The mean square error $v^{(k)}_n := E(X_{t+k} - \hat{X}^{(k)}_{t+k})^2$ of the best k-step linear prediction is computed by the formula

$$(4.3) \qquad \boxed{v^{(k)}_n = \gamma(0) - \sum_{j=1}^{n} \phi^{(k)}_{nj} \gamma(k+j-1) = \gamma(0) - \underline{\phi}^T_n \cdot \underline{\gamma}^{(k)}_n}$$

Proof:

$$v^{(k)}_n = E(X_{t+k} - \hat{X}^{(k)}_{t+k})^2 = \| X_{t+k} - \hat{X}^{(k)}_{t+k} \|^2 = \langle X_{t+k} - \hat{X}^{(k)}_{t+k}, X_{t+k} - \hat{X}^{(k)}_{t+k} \rangle =$$

$$= \langle X_{t+k} - \hat{X}^{(k)}_{t+k}, X_{t+k} \rangle - \underbrace{\langle X_{t+k} - \hat{X}^{(k)}_{t+k}, \hat{X}^{(k)}_{t+k} \rangle}_{=0 \text{ because } X_{t+k} - \hat{X}^{(k)}_{t+k} \perp \hat{X}^{(k)}_{t+k}} = \langle X_{t+k}, X_{t+k} \rangle - \langle \underbrace{\hat{X}^{(k)}_{t+k}}_{\sum_{j=1}^{n} \phi^{(k)}_{nj} X_{t+1-j}}, X_{t+k} \rangle =$$

$$= \underbrace{\| X_{t+k} \|^2}_{E X^2_{t+k} = \sigma^2_X = \gamma(0)} - \sum_{j=1}^{n} \phi^{(k)}_{nj} \underbrace{\langle X_{t+1-j}, X_{t+k} \rangle}_{\operatorname{cov}(X_{t+1-j}, X_{t+k}) = \gamma(t+1-j-(t+k)) = \gamma(-(k+j-1)) \overset{\text{by sym.}}{=} \gamma(k+j-1)}$$