Stochastic Processes

w.s.s.

$$\mathrm{E}[X[n]]=\text{const.}\,,\mathrm{Var}(X[n])=\text{const.}<\infty$$

$$\mathrm{E}[X[k]X^*[l]]=R(k-l)\,,\forall k,l\in\mathbb{Z}\,,$$

PSD: w.s.s. + $R_X(k)\in\ell_1$ (summable)

$$S_X(\omega)=\sum_{k=-\infty}^{\infty}R_X(k)e^{-i\omega k}\,.$$

Fundamental Filtering Formula

$X[n]$ w.s.s., $R_X(k)\in\ell_1$, and $h_k\in\ell_1$, then

$$Y[n]=\sum_{k=-\infty}^{\infty}h_{n-k}X[k]\,,\text{ is w.s.s. with}$$

$$\mathrm{E}[Y]=\mathrm{E}[X]\sum_{k=-\infty}^{\infty}h_k\,,\quad R_Y(k)\in\ell_1$$

$$S_Y(\omega)=|H(\omega)|^2S_X(\omega)\,,H(\omega)=\text{DTFT of }h_k$$

Markov Chain

$\{X[n]\}_{n\in\mathbb{Z}}$ (considered stationary) with discrete values in $\mathcal{D}$, |

$$\mathrm{P}\big(X[n]=i_n\mid X[n-1]=i_{n-1},X[n-2]=i_{n-2},\ldots\big)$$
$$=\mathrm{P}\big(X[n]=i_n\mid X[n-1]=i_{n-1}\big)\,,\forall i_n,i_{n-1},\ldots\in\mathcal{D}$$

Hidden Markov Chain

$\{X[n]\}_{n\in\mathbb{Z}}$ Markov chain, $\{W[n]\}_{n\in\mathbb{Z}}$ Gaussian white noise

$$Y[n]=X[n]+W[n]\,.$$

Bayes' Rule

A and B with discrete values in $\mathcal{D}$,

$$\mathrm{P}(A=k\mid B=l)=\frac{\mathrm{P}(A=k\,,\,B=l)}{\mathrm{P}(B=l)}\,,\quad k,l\in\mathcal{D}\,.$$

AR Process

A w.s.s. process $X[n]$, with values in $\mathbb{R}$, |

$$\sum_{k=0}^{M}p_kX[n-k]=W[n]\,,\quad n\in\mathbb{Z}\,,$$

$W[n]$, is a zero mean Gaussian white noise

$p_k\,,k=0,\ldots,M$ bounded coefficients (real or complex). We assume $p_0=1$.

Filtering Interpretation (z^{-1} delay operator)

$$P(z)X[n]=W[n]$$

Canonical form: $P(z)$ strict. min. phase, $p_0=1$.

Correlation:

$$R_X[m]+\sum_{k=1}^{M-1}p_kR_X[m-k]=\delta_m\sigma_W^2\,,\quad m\geq 0\,.$$

PSD (fundamental filtering formula):

$$S_X(\omega)\big|P\big(e^{j\omega}\big)\big|^2=\sigma_W^2\,.$$

Harmonic Processes

$$X[n]=\sum_{k=1}^{K}\alpha_ke^{j(\omega_kn+\Theta_k)}\,,n\in\mathbb{N}\,,$$

Θ_k i.i.d. uniformly distributed over $[0,2\pi]$.

$$R_X[l]=\sum_{k=1}^{K}|\alpha_k|^2e^{j\omega_kl}\,,\;S_X(\omega)=\sum_{k=1}^{K}|\alpha_k|^2\delta(\omega-\omega_k)\,.$$

Poisson Random Process $N((0,t])$

$N((0,t])$ obeys the Poisson distribution $\mathrm{P}(N=k)=\frac{(a\lambda)^ke^{-a\lambda}}{k!}$, ($\lambda$ is the rate), and given two disjoint intervals $(t_1,t_2]$ and $(t_3,t_4]$, $N((t_1,t_2])$ is independent of $N((t_3,t_4])$.

Inter-arrival time $S_n=T_n-T_{n-1}$. i.i.d. with density $f_S(t)=\lambda e^{-t\lambda}$.

Hilbert Spaces

Projection Theorem

E,S Hilbert spaces with $S\subset E$, then

$$\forall v\in E\,,\exists!\,b\in S\mid$$

$$b=\arg\min_{c\in S}\|v-c\|\,,\;\langle v-b,c\rangle=0\,,\forall\;c\in S\,,$$

Projection Theorem w.s.s.

$E\;S$ Hilbert spaces of w.s.s. processes with $S\subset E\subseteq \mathrm{L}^2(P)$, then

$$\forall X[n]\in E\,,\exists!\,Y[n]\in S\mid$$

$$Y[n]=\arg\min_{U[n]\in S}\mathrm{E}[|X[n]-U[n]|^2]\,,$$

$$\mathrm{E}[(X[n]-Y[n])U^*[n]]=0\,,\forall\;U[n]\in S\,,$$

Empirical Statistics

Bias & Variance

$\widehat{S}(x[1],\ldots,x[N])$ empirical statistics of a probabilistic moment S.

Bias $\mathrm{E}\big[\widehat{S}(X[1],\ldots,X[N])\big]-S$,

Variance $\mathrm{Var}\big(\widehat{S}(X[1],\ldots,X[N])-S\big)$.

Unbiased & Biased Correlation

$$\widehat{R}_X^{\mathrm{NB}}(k)=\frac{1}{N-|k|}\sum_{n=1}^{N-|k|}x[n+k]x^*[n]\,,$$

$$\widehat{R}_X^{\mathrm{B}}(k)=\frac{1}{N}\sum_{n=1}^{N-|k|}x[n+k]x^*[n]\,.$$

Methods

Linear Estimation of w.s.s.: Wiener Filter

Estimation of $X[n]$ given $Y[n]$

Normal equations $R_{XY}[u]=\sum_{m\in\mathbb{Z}}h[m]R_Y[u-m]$

Wiener Filter $H(e^{j\omega})=\frac{S_{XY}(\omega)}{S_Y(\omega)}$

Linear Prediction of w.s.s.: Yule-Walker

Prediction of $X[n]$ as linear combination of $X[n-1],\ldots,X[n-N]$.

Coefficients a_k solutions of

$$\begin{bmatrix}R_X[0]&\ldots&R_X[N-1]\\\vdots&\ddots&\vdots\\R_X[N-1]&\ldots&R_X[0]\end{bmatrix}\begin{bmatrix}a_1\\\vdots\\a_N\end{bmatrix}=\begin{bmatrix}R_X[1]\\\vdots\\R_X[N]\end{bmatrix}.$$

Linear Estimation of AR: Yule-Walker

$\sum_{k=0}^{N}p_kX[n-k]=W[n]$. Coeff. p_k solution of

$$-\begin{bmatrix}R_X[0]&\ldots&R_X[N-1]\\\vdots&\ddots&\vdots\\R_X[N-1]&\ldots&R_X[0]\end{bmatrix}\begin{bmatrix}p_1\\\vdots\\p_N\end{bmatrix}=\begin{bmatrix}R_X[1]\\\vdots\\R_X[N]\end{bmatrix}.$$

$$\sigma_W^2=R_X[0]+R_X[1]p_1+\ldots+R_X[N]p_N$$

Linear Pretiction of AR: Projection Theorem

$H(X,n)=H(W,n)\,,\quad\forall\;n\in\mathbb{Z}\,.$

Intuitive property:

$$Y\in H(X,n+k)\,,\;Y=A+B\,,\;A\perp H(X,n)\,,\;B\in H(X,n)$$

orthogonal projection of Y onto $H(X,n)$ is B.

Estimation Param. Prob.: MLE

$\boldsymbol{\theta}$ parameters of the prob. function f_Y, $y[1],\ldots,y[N]$ realization of the process $Y[n]$, then

$$\widehat{\boldsymbol{\theta}}=\arg\max_{\boldsymbol{\theta}}f_Y(y[1],\ldots,y[N],\boldsymbol{\theta})$$

Spectral Estimation

Periodogram: General w.s.s. process

$$P_X^N(\omega)=\frac{1}{N}\left|\sum_{n=1}^{N}x[n]e^{-j\omega n}\right|^2=\frac{1}{N}|\widehat{x}_N(\omega)|^2\,,$$

Bias $\sum_{k=-N+1}^{N-1}\frac{N-|k|}{N}R_X[k]e^{-j\omega k}-S_X(\omega)$

Variance constant

Resolution $\Delta f>\frac{1}{N}$

Annihilating Filter: Line Spectra

Estimation of line spectrum frequencies and amplitudes of a Harmonic w.s.s. process in absence of noise

1) Given $2K$ observations, solve the system

$$\begin{bmatrix}x[K-1]&\ldots&x[0]\\\vdots&&\vdots\\x[2K-2]&\ldots&x[K-1]\end{bmatrix}\begin{bmatrix}h[1]\\\vdots\\h[K]\end{bmatrix}=-\begin{bmatrix}x[K]\\\vdots\\x[2K-1]\end{bmatrix}$$

2) Compute $H(z)$ and the zeros of $H(z)$

3) Compute the argument of the zeros of $H(z)$

4) Compute ω_k from the zeros' arguments

5) Compute the amplitudes $|\alpha_k|^2$ by solving

$$\begin{bmatrix}1&\ldots&1\\e^{j\omega_1}&\ldots&e^{j\omega_K}\\\vdots&&\vdots\\e^{j\omega_1(K-1)}&\ldots&e^{j\omega_K(K-1)}\end{bmatrix}\begin{bmatrix}\alpha_1e^{j\Theta_1}\\\alpha_2e^{j\Theta_2}\\\vdots\\\alpha_Ke^{j\Theta_K}\end{bmatrix}=\begin{bmatrix}x[0]\\x[1]\\\vdots\\x[K-1]\end{bmatrix}$$

MUSIC: Line Spectra

Estimation of line spectrum frequencies and amplitudes of a Harmonic w.s.s. process in the presence of noise

1) Given M observations with $M>>N>K$ center the process and compute the empirical correlation matrix

$$\widehat{\boldsymbol{R}}_{\boldsymbol{Y}}^{NN}=\frac{1}{M-N+1}\sum_{n=1}^{M-N+1}\boldsymbol{y}^{N1}[n]\boldsymbol{y}^{N1}[n]^H\,;$$

2) Compute the eigendecomposition $\widehat{\boldsymbol{G}}^{N(N-K)}$ of $\widehat{\boldsymbol{R}}_{\boldsymbol{Y}}^{NN}$ corresponding to λ_{K+1}^R to λ_N^R.

3.a) Determine the peaks of

$$\frac{1}{\boldsymbol{e}^{N1}(\omega)^H\widehat{\boldsymbol{G}}^{N(N-K)}\widehat{\boldsymbol{G}}^{N(N-K)\,H}\boldsymbol{e}^{N1}(\omega)}\,;$$

where

$$\boldsymbol{e}^{N1}(\omega)=\begin{bmatrix}1&e^{-j\omega}&\ldots&e^{-j(N-1)\omega}\end{bmatrix}^T,$$

3.b) Determine the minimum values of

$$\boldsymbol{e}^{N1}(\omega)^H\widehat{\boldsymbol{G}}^{N(N-K)}\widehat{\boldsymbol{G}}^{N(N-K)\,H}\boldsymbol{e}^{N1}(\omega)\,.$$

4) Compute the modulus of the amplitudes using

$$\boldsymbol{R}_{\boldsymbol{Y}}^{NN}=\boldsymbol{E}^{NK}\boldsymbol{A}^{KK}\boldsymbol{E}^{NK\,H}+\sigma_W^2\boldsymbol{I}^{NN}\,.$$

Yule-Walker: Smooth Spectra

1) Given N observations with $N>>M$ center the process and

compute the empirical correlation

$$\widehat{R}_X[k]=\frac{1}{N}\sum_{n=1}^{N-k} x_0[n+k]x_0[n]^*\,, \quad k=0,\dots,M$$

2) Solve the Yule Walker equations to obtain $\hat{p}_1,\dots,\hat{p}_M$

3) Compute the estimate of the spectrum as

$$\widehat{S}_X(\omega)=\left.\frac{\widehat{\sigma}_W^2}{|\widehat{P}(z)|^2}\right|_{z=e^{j\omega}}.$$

Mixture Models

Sequence of samples $\boldsymbol{y}=[y[1],\dots,y[N]]$, sequence of corresponding classes $\boldsymbol{c}=[c[1],\dots,c[N]]$

$$F_{\boldsymbol{Y}}(\boldsymbol{y})=\sum_{\boldsymbol{c}\in\mathcal{C}} \mathrm{P}(\boldsymbol{Y}\le\boldsymbol{y},\boldsymbol{C}=\boldsymbol{c})=\sum_{\boldsymbol{c}\in\mathcal{C}} \mathrm{P}(\boldsymbol{Y}\le\boldsymbol{y}\mid\boldsymbol{C}=\boldsymbol{c})\mathrm{P}(\boldsymbol{C}=\boldsymbol{c})$$

$$=\sum_{\boldsymbol{c}\in\mathcal{C}}\prod_{n=1}^{N} \mathrm{P}(Y[n]\le y[n]\mid C[n]=c[n])\mathrm{P}(\boldsymbol{C}=\boldsymbol{c})$$

i.i.d. Mixtures

$$\mathrm{P}(\boldsymbol{C}=\boldsymbol{c})=\mathrm{P}(C[1]=c[1])\dots\mathrm{P}(C[N]=c[N])=\pi_{c[1]}\cdots\pi_{c[N]}\,,$$

Markovian Mixtures

$$\mathrm{P}(\boldsymbol{C}=\boldsymbol{c})=\pi_{c[1]}p_{c[1]c[2]}\cdots p_{c[N-1]c[N]}\,,$$

Hidden Markov chain

$\boldsymbol{Y}=\boldsymbol{X}+\boldsymbol{W}$ where $\boldsymbol{X}$ is a Markov Chain and $\boldsymbol{W}$ a white Gaussian noise.

$$f_{\boldsymbol{Y}}(\boldsymbol{y})=\sum_{\boldsymbol{x}\in\mathcal{C}}\prod_{n=1}^{N}\mathcal{G}_{x[n],\sigma^2}(y[n])\mathrm{P}(\boldsymbol{X}=\boldsymbol{x})\,.$$

Denoising a Discrete Value Process

Estimate the parameters of the mixture model using the maximum likelihood approach; Estimate the original signal using the maximum *a posteriori* approach, *i.e.*, find $\boldsymbol{x}$ maximizing the *a posteriori* distribution

$$\mathrm{P}(\boldsymbol{X}=\boldsymbol{x}\mid\boldsymbol{y})=\frac{f_{\boldsymbol{Y}}(\boldsymbol{y}\mid\boldsymbol{X}=\boldsymbol{x})\mathrm{P}(\boldsymbol{X}=\boldsymbol{x})}{f_{\boldsymbol{Y}}(\boldsymbol{y})}$$

PCA

Principal Components Computaiton

M data vectors, each characterized of N variables (realization of a zero mean w.s.s. process) $\boldsymbol{c}_m=[c_m[1],\dots,c_m[N]]^T\,,\quad m=1,\dots,M.$

Empirical correlation matrix

$\widehat{\boldsymbol{R}}_{\boldsymbol{c}}=\frac{1}{M}\sum_{m=1}^{M}\boldsymbol{c}_m*\boldsymbol{c}_m^H=\frac{1}{M}\boldsymbol{C}*\boldsymbol{C}^H\,,(N\times N)\,,$ where $\boldsymbol{C}=\begin{bmatrix}\boldsymbol{c}_1 & \dots & \boldsymbol{c}_M\end{bmatrix},$

V solution of the equation $\widehat{\boldsymbol{R}}_{\boldsymbol{c}}\boldsymbol{V}=\boldsymbol{V}\boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\dots,\lambda_N)$ and $\boldsymbol{V}^H\widehat{\boldsymbol{R}}_{\boldsymbol{c}}\boldsymbol{V}=\boldsymbol{\Lambda}$.

Principal components $\boldsymbol{Z}=\boldsymbol{V}^T\boldsymbol{C}\,,(N\times M)$, uncorrelated.

Invertible transformation $\boldsymbol{C}=\boldsymbol{V}\boldsymbol{Z}\,,$

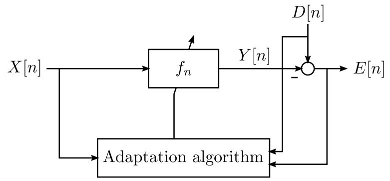
Analysis

$K<<N$ eigenvalues with highest values (lossy/lossless reduction of variables)

Adaptive Filtering / Echo cancellation

Wiener-Hopf equations $\sum_{k\in\mathbb{Z}} h[n-k]R_Y(k-l)=R_{XY}(n-l)\,,\quad\forall\, l\,.$

Echo cancellation setup



$$E[n]=D[n]-f_n*X[n]=S[n]+h*X[n]-f_n*X[n]=S[n]+(h-f_n)*X[n]\,.$$

Cost function & normal equations

Cost function for a k-tap filter

$$J(f_n)=\mathrm{E}\Big[|E[n]|^2\Big]=\mathrm{E}\Big[|D[n]-f_n*X[n]|^2\Big]\,,\quad \text{w.r.t.}\quad f_n[l]\,,l=0,1,\dots,K$$

Minimum of the cost function = normal equations

$$\sum_{l=0}^{K} f_n[l]R_X(n-l,n-i)=R_{DX}(n,n-i)\,,\quad \boldsymbol{R}_{X,n}\boldsymbol{f}_n=\boldsymbol{r}_{DX,n}\,.$$

Iterative solution

$$\boldsymbol{f}^{(i+1)}=\boldsymbol{f}^{(i)}+\mu\boldsymbol{p}\,,\quad i=0,1,\dots,\quad \mu\ \&\ \boldsymbol{p}\text{ such that } J(\boldsymbol{f}^{(i+1)})<J(\boldsymbol{f}^{(i)})$$

Convergence conditions

$$0<\mu<2/\lambda_{\max}\,,\quad \boldsymbol{p}=\tfrac{1}{2}(\boldsymbol{r}_{DX}-\boldsymbol{R}_X\boldsymbol{f}^{(i)})\text{ or }\boldsymbol{p}=4\boldsymbol{R}_X^{-1}(\boldsymbol{r}_{DX}-\boldsymbol{R}_X\boldsymbol{f}^{(i)})(\text{Newton})$$

Convergence rate

- $0\le 1-\mu\lambda_j<1$, monotonic decay to zero;
- $-1<1-\mu\lambda_j<0$ oscillatory decay to zero.
- $K+1$ modes $\{1-\mu\lambda_j, j=0,\dots,K\}$. The modes with maximum magnitude (slowest rate of convergence), determine the convergence rate of the algorithm. One can select μ optimally by minimizing the value of the slowest mode $\min_\mu\max_{j=0,\dots,K}|1-\mu\lambda_j|$, with the constraint that each of the modes is stable, i.e., $|1-\mu\lambda|<1$.

Computational burden reduction

Merging interation and adaptation

$$\boldsymbol{f}_{n+1}=\boldsymbol{f}_n+\mu(\boldsymbol{r}_{DX,n+1}-\boldsymbol{R}_{X,n+1}\boldsymbol{f}_n)\,,$$

$$\boldsymbol{f}_{n+1}=\boldsymbol{f}_n+\mu\boldsymbol{R}_{X,n+1}^{-1}(\boldsymbol{r}_{DX,n+1}-\boldsymbol{R}_{X,n+1}\boldsymbol{f}_n)\quad(\text{Newton})\,,$$

Replacing statistics with individual values

$$\boldsymbol{f}_{n+1}=\boldsymbol{f}_n+\mu\boldsymbol{X}_n(D[n]-\boldsymbol{X}_n^T\boldsymbol{f}_n)=\boldsymbol{f}_n+\mu\boldsymbol{X}_nE[n]\,,$$