Auto-regressive model

$$\hat{y}[n] = a_1 y[n-1] + a_2 y[n-2] + \cdots + a_p y[n-p] + \underbrace{\epsilon[n]}_{\text{רעש}}$$
$$= \sum_{m=1}^{p} a_m y[n-m] + \epsilon[n]$$

p - hyper-parameter

$a_1, \ldots, a_p$ - פרמטרים של מודל הניבוי

$$\underbrace{\begin{bmatrix} \hat{y}[1] \\ \hat{y}[2] \\ \vdots \\ \hat{y}[L-1] \\ \hat{y}[L] \end{bmatrix}}_{\hat{\mathbf{y}}} = \underbrace{\begin{bmatrix} y[0] & 0 & 0 & \vdots \\ y[1] & y[0] & 0 & \\ y[2] & y[1] & y[0] & \\ y[3] & y[2] & y[1] & \\ \vdots & \vdots & \vdots & \vdots \\ y[L-2] & y[L-3] & y[L-2] & \\ y[L-1] & y[L-2] & y[L-3] & \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}}_{\mathbf{a}},$$

$= a_1 y[0]$

$a_1 y[1] + a_0 y[0]$

p שת

L-1 שורות

L-1 × p ↓ שורות

פיני מחיר :

$$\mathcal{L}(a_i) = \sum_n \left(y[n] - \hat{y}[n]\right)^2 = \left\| \mathbf{y} - \mathbf{X}\bar{\hat{\mathbf{y}}} \right\|^2$$

$$\mathbf{a} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1} \mathbf{X}^T\mathbf{y}$$

Auto-correlation matrix (Toeplitz)

$p << L$

$$\mathbf{X}^T\mathbf{X} = \begin{bmatrix} R_{\mathbf{yy}}[0] & R_{\mathbf{yy}}[1] & \cdots & R_{\mathbf{yy}}[p-1] \\ R_{\mathbf{yy}}[1] & R_{\mathbf{yy}}[0] & \cdots & R_{\mathbf{yy}}[p-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_{\mathbf{yy}}[p-1] & R_{\mathbf{yy}}[p-2] & \cdots & R_{\mathbf{yy}}[0] \end{bmatrix}$$

$$R_{\mathbf{yy}}[0] = \sum_{n=0}^{L-1} y^2[n] \approx \sum_{n=0}^{L-2} y^2[n]$$

$$R_{\mathbf{xx}}[k] = \sum_n x[n]x[n-k], \quad k = 0, \ldots, L-1$$

$$y = \begin{bmatrix} y[1] \\ \vdots \\ y[L-1] \end{bmatrix}, \quad X = \underbrace{\begin{bmatrix} y[0] & 0 & 0 & \vdots \\ y[1] & y[0] & 0 & \\ y[2] & y[1] & y[0] & \\ y[3] & y[2] & y[1] & \\ \vdots & \vdots & \vdots & \vdots \\ y[L-2] & y[L-3] & y[L-2] & \\ y[L-1] & y[L-2] & y[L-3] & \end{bmatrix}}_{\mathbf{X}}$$

$$\mathbf{X}^T\mathbf{y} = \begin{bmatrix} R_{\mathbf{yy}}[1] \\ R_{\mathbf{yy}}[2] \\ \vdots \\ R_{\mathbf{yy}}[p+1] \end{bmatrix}$$

Yule-Walker algorithm ⟸ חלופה מהירה ⟸ $\mathbf{a} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$

$$\mathcal{L}_{min} = \sum_{n=0}^{L-1} x^2[n] - \sum_{k=1}^{p} a_k \sum_{n=0}^{L-1} x[n]x[n-k]$$
$$= R_{\mathbf{xx}}[0] - \sum_{k=1}^{p} a_k R_{\mathbf{xx}}[k]$$

ACF = מדד לקשר ליניארי בזמן

Example 18.1: Learn linear prediction of $y[8]$ for $p = 2$ and signal $y[0], y[1], \ldots, y[7]$.

Solution:

$$\underbrace{\begin{bmatrix} y[0] & 0 & 0 \\ y[1] & y[0] & 0 \\ y[2] & y[1] & y[0] \\ y[3] & y[2] & y[1] \\ y[4] & y[3] & y[2] \\ y[5] & y[4] & y[3] \\ y[6] & y[5] & y[4] \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}}_{\mathbf{a}} = \underbrace{\begin{bmatrix} y[1] \\ y[2] \\ y[3] \\ y[4] \\ y[5] \\ y[6] \\ y[7] \end{bmatrix}}_{\mathbf{y}} \tag{18.9}$$

Finding vector **a** values is (almost) trivial by a minimum of the loss function of the form

$$\mathcal{L} = \|\mathbf{y} - \mathbf{X}\mathbf{a}\|^2. \tag{18.10}$$

Finally, $\hat{y}[8] = a_1 y[7] + a_2 y[6] + a_3 y[5]$.

Biased signal When the signal of interest in non-zero average, the biased model of the form is used,

$$\hat{y}[n] = \mu + a_1 y[n-1] + a_2 y[n-2] + \cdots + a_p y[n-p] + \epsilon[n] \tag{18.11}$$

DC רכיב

$$\hat{y}[n] = \mu + a_1 y[n-1] + a_2 y[n-2] + \cdots + a_p y[n-p] + \epsilon[n] \quad (18.11)$$

רכיב DC

פתרון: הוספת עמודת אחדות בתוך X

$$\begin{bmatrix} \mathbf{1}_{L-1} & \mathbf{x}_1 & \cdots & \mathbf{x}_N \end{bmatrix}$$

L-1 אחדות

Finally, $\hat{y}[8] = a_1 y[7] + a_2 y[6] + a_3 y[5]$.

Confidence Interval Another way to quantify the "importance" of the coefficients is 95% confidence bound is given by $\Delta = \pm \frac{\sqrt{2}\,\mathrm{erf}(0.95)}{\sqrt{L}} = \pm \frac{1.96}{\sqrt{L}}$

ערכים "חשובים" של ACF

Filter command in Matlab

$a_0 = 1$

$$a(1)y(L) = b(1)x(L) + b(2)x(L-1) + \ldots + b(n_b+1)x(L-n_b)$$
$$-\underbrace{a(2)}_{a_1}y(L-1) - \ldots - \underbrace{a(n_a+1)}_{a_p}y(L-n_a).$$

Partial auto-correlation function

$$x[0], \underbrace{x[1], x[2], \ldots, x[j-1]}_{\text{partial out}}, x[j], x[j+1], \ldots$$

מטרה: "לבטל" השפעה של המקדמים "הקרובים" של חיזוי

מספר ערכי K משמעותיים (פרט לK=0) הוא בקירוב מספר p הנדרש של חיזוי

$K=0 \quad \beta[0] = 1$

$K=1 \quad \hat{y}_1[n] = \underbrace{\phi_{1,1}}_{\beta[1]} y[n-1] + \epsilon[n]$

$K=2 \quad \hat{y}_2[n] = \phi_{2,1} y[n-1] + \underbrace{\phi_{2,2}}_{\beta[2]} y[n-2] + \epsilon[n]$

$K=3 \quad \phi_{3,3}$