FIR

$\hat{y}(t|\theta) = \theta^T \varphi(t)$

$y(t) = \sum_{i=1}^{nb} b_i \, u(t-i) + e(t) = \theta^T \varphi(t) + e(t)$

$\hat{y}(t) = \theta^T \varphi(t)$

$\varepsilon(t,\theta) = y(t) - \hat{y}(t|t-1,\theta)$

$\hat{\theta}_N = \arg\min_{\theta} V_N(\theta)$

- $V_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} \varepsilon^2(t,\theta)$
- $V_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} \left(y(t) - \hat{y}(t|t-1,\theta)\right)^2$

$$\nabla_\theta V_N(\theta) = -\frac{2}{N} \sum \underbrace{\nabla_\theta \hat{y}(t|t-1,\theta)}_{\varphi} \left(y(t) - \hat{y}(t|t-1,\theta)\right) = 0$$

$$\left[\frac{1}{N} \sum_{t=1}^{N} \varphi(t) \varphi^T(t)\right] \cdot \theta = \frac{1}{N} \sum_{t=1}^{N} \varphi(t) \cdot y(t)$$

→ this system of linear equations gives the solution to $\hat{\theta}$ as before in the least squares case:

$$\hat{\theta} = R_N^{-1} f_N$$

OE :

$$\hat{y}(t) = (1 - F(q))\hat{y}(t) + B(q).u(t) = \theta^T.\varphi(t,\theta)$$

$$\theta^T = [f_1 \cdots f_{nf}, b_1 \cdots b_{n_b}]$$

$$\varphi(t,\theta) = [-\hat{y}(t-1) \cdots -\hat{y}(t-n_f), u(t-1) \cdots u(t-n_b)]$$

$$\frac{\partial}{\partial b_i}\hat{y}(t) = (1-F(q))\frac{\partial \hat{y}(t)}{\partial b_i} + u(t-i)$$

$$\frac{\partial}{\partial f_i}\hat{y}(t) = (1-F(q))\frac{\partial \hat{y}(t)}{\partial f_i} - \hat{y}(t-i)$$

$$F\frac{\partial \hat{y}}{\partial b} = u \qquad F\begin{bmatrix} \frac{\partial \hat{y}}{\partial f} \\ \frac{\partial \hat{y}}{\partial b} \end{bmatrix} = \begin{bmatrix} -y \\ u \end{bmatrix}$$

$$F\frac{\partial \hat{y}}{\partial f} = -\hat{y}$$

$$F\,\nabla_\theta \hat{y}(t) = \varphi(t,\theta)$$

$$\nabla_\theta \hat{y}(t) = \frac{1}{F(q,\theta)}\varphi(t,\theta)$$

$\hookrightarrow$ we need this gradient to find the $\hat{\theta}$ parameter estimate while optimizing the cost function.