

16 & 17.

$$\hat{\theta} = \Phi_N^+ y_N$$

$$(\Phi_N^T \Phi_N)^{-1} \Phi_N^T = \Phi_N^+$$

$$y_N = \Phi_N \theta_0 + \varepsilon_N$$

$$\hat{\theta} = \Phi_N^+ (\Phi_N \theta_0 + \varepsilon_N) = \theta_0 + \underbrace{\Phi_N^+ \varepsilon_N}_{\equiv}$$

$$\Sigma_{\hat{\theta}} = (\Phi_N^+)^T \left(\Sigma_{\varepsilon_N} \right) \Phi_N^+$$

16: $\Sigma_{\varepsilon_N} = \sigma^2 I$

$$\Rightarrow \Sigma_{\hat{\theta}} = \sigma^2 (\Phi_N^T \Phi_N)^{-1}$$



17.

$$\left[W = \Sigma_{\varepsilon_N}^{-1} \right]$$

$$\Rightarrow \hat{\theta} = \underset{\theta}{\text{argmin}} \| y_N - \Phi_N \theta \|_W^2$$

$$\Sigma_{\hat{\theta}} = (\Phi_N^T W \Phi_N)^{-1}$$

18. $\theta = 40\%$??

$$\|x\|_W^2 = x^T W x$$

$$p(\text{heads} | \theta) = \theta$$

40 x

$$p(\text{tails} | \theta) = 1 - \theta$$

60 x

$$y_N = \begin{bmatrix} \text{heads} \\ \vdots \\ \text{heads} \\ \text{tails} \\ \vdots \\ \text{tails} \end{bmatrix}$$

$$L(\theta) = p(y_N | \theta) = \prod_{i=1}^{100} p(y_i | \theta) = \prod_{i=1}^{40} \theta \cdot \prod_{i=1}^{60} (1 - \theta)$$

$$\begin{aligned}
 -\log L(\theta) &= \sum_{i=1}^{40} -\log \theta + \sum_{i=1}^{60} -\log(1-\theta) \\
 &= 40 \log \theta - 60 \log(1-\theta) \rightarrow C \\
 &=: f(\theta)
 \end{aligned}$$

$$\frac{\partial f}{\partial \theta} = \frac{-40}{\theta} + \frac{60}{1-\theta} \stackrel{!}{=} 0$$

$$\Leftrightarrow \frac{40}{\theta} = \frac{60}{1-\theta} \Leftrightarrow \frac{40}{60} = \frac{\theta}{1-\theta}$$

$$\boxed{\theta = 0.4}$$

