1. Determine the MAP estimate $\widehat{\theta}_{MAP}(\mathbf{x})$ **given a scaled signal in noise.**

$$\mathbf{x} = \theta\mathbf{s} + \mathbf{w}\,, \qquad \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma_w^2\mathbf{I})\,, \qquad p(\theta) = \begin{cases} \lambda\exp(-\lambda\theta) & \text{for } \theta \geq 0 \\ 0 & \text{for } \theta < 0 \end{cases}$$

As shown in lecture *33 - MAP Estimation*, the maximum a posteriori estimator is that which maximizes the a posteriori probability estimate (what a suprise!) which visually corresponds to the peak of the resulting a posteriori probability density function (PDF).

$$\widehat{\theta}_{MAP}(\mathbf{x}) = \arg\left[\max_{\widehat{\theta}}\left\{p(\widehat{\theta}|\mathbf{x})\right\}\right]$$

The a posteriori density function can be related to the a priori PDF, measurement PDF, and liklihood function through Bayes theorem and, that the denominator is independent of $\widehat{\theta}$ and we do not need to find the actual value of the maximum only the $\widehat{\theta}$ it corresponds to, the denominator can be ignored leaving

$$\widehat{\theta}_{MAP}(\mathbf{x}) = \arg\left[\max_{\widehat{\theta}}\left\{p(\mathbf{x}|\theta)p(\theta)\right\}\right].$$

Given that the additive noise is independent and identically distributed normally, the liklihood function will likewise be normal with mean $\theta\mathbf{s}$ and variance σ_w^2.

$$p(\mathbf{x}|\theta) = (2\pi\sigma_w^2)^{-N/2}\exp\left[-\frac{|\mathbf{x}-\theta\mathbf{s}|^2}{2\sigma_w^2}\right]$$

Again, being that we only need to know the value of $\widehat{\theta}$ that maximizes the PDF not the max of the PDF itself, the natural logarithm, being monotonically increasing, is applied to argument of the $\max\{\,*\,\}$ operator

$$\ln\left(p(\mathbf{x}|\theta)p(\theta)\right) = \ln\left(p(\mathbf{x}|\theta)\right) + \ln(p(\theta))$$

where

$$\ln(p(\mathbf{x}|\theta)) = -\frac{N}{2}\ln(2\pi\sigma_w^2) - \frac{\mathbf{x}^\mathbf{T}\mathbf{x} - 2\mathbf{s}^\mathbf{T}\mathbf{x}\theta + \mathbf{s}^\mathbf{T}\mathbf{s}\theta^2}{2\sigma_w^2}, \qquad \ln(p(\theta)) = \ln(\lambda) - \lambda\theta \;\text{ for } \theta \geq 0 \;\text{ else }\; 0$$

therefore

$$\ln\left(p(\mathbf{x}|\theta)p(\theta)\right) = -\frac{N}{2}\ln(2\pi\sigma_w^2) - \frac{\mathbf{x}^\mathbf{T}\mathbf{x} - 2\mathbf{s}^\mathbf{T}\mathbf{x}\theta + \mathbf{s}^\mathbf{T}\mathbf{s}\theta^2}{2\sigma_w^2} + \ln(\lambda) - \lambda\theta$$

$$= -\frac{\mathbf{s}^\mathbf{T}\mathbf{s}}{2\sigma_w^2}\theta^2 + \left(\frac{\mathbf{s}^\mathbf{T}\mathbf{x}}{\sigma_w^2} - \lambda\right)\theta - \frac{N}{2}\ln(2\pi\sigma_w^2) - \frac{\mathbf{x}^\mathbf{T}\mathbf{x}}{2\sigma_w^2} + \ln(\lambda).$$

The above term is identified to be quadratic and can easily be differentiated to determine the extreme points and check for convexity/concavity to find the maximum. The first derivative is taken below

$$\frac{\partial \ln(p(\mathbf{x}|\theta)p(\theta))}{\partial\theta} = -\frac{\mathbf{s}^{\mathbf{T}}\mathbf{s}}{\sigma_w^2}\theta + \frac{\mathbf{s}^{\mathbf{T}}\mathbf{x}}{\sigma_w^2} - \lambda .$$

Setting this equal to zero gives the extreme point

$$-\frac{\mathbf{s}^{\mathbf{T}}\mathbf{s}}{\sigma_w^2}\theta + \frac{\mathbf{s}^{\mathbf{T}}\mathbf{x}}{\sigma_w^2} - \lambda = 0 \quad \rightarrow \quad \widehat{\theta}_{MAP} = \frac{\mathbf{s}^{\mathbf{T}}\mathbf{x} - \lambda\sigma_w^2}{\mathbf{s}^{\mathbf{T}}\mathbf{s}}.$$

The second derivative is applied to determine concavity and prove that this extreme point is a maximum

$$\frac{\partial^2 \ln(p(\mathbf{x}|\theta)p(\theta))}{\partial^2\theta} = -\frac{\mathbf{s}^{\mathbf{T}}\mathbf{s}}{\sigma_w^2}.$$

Knowing that $\sigma_w^2 \geq 0$ and $\mathbf{s}^{\mathbf{T}}\mathbf{s} = \mathbf{s}^2 \geq 0$, this result will always be negative therefore $\widehat{\theta}_{MAP}$ proven to be a maximum. Summarizing the results derived above, the estimate of θ that maximizes the a posteriori PDF was found as

$$\widehat{\theta}_{MAP} = \frac{\mathbf{s}^{\mathbf{T}}\mathbf{x} - \lambda\sigma_w^2}{\mathbf{s}^{\mathbf{T}}\mathbf{s}}$$

and proven to be a maximum using the second derivative test for concavity. This result can be interpreted as, in the numerator, filtering the measurement $\mathbf{x}$ with the 'matched' known signal vector $\mathbf{s}$, then regularizing by the noise variance and the priori strength of source θ. The denominator then normalizes this result by the unscaled signal energy.