1.9 First we show that the mode of the uni-variate Gaussian distribution is given by μ.

The distribution is given by:

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{2\pi\sigma^2} exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}$$

The mode is the maximum, so we can find it by taking the derivative of the function w.r.t x and setting it to 0:

$$\frac{d}{dx}\mathcal{N}(x|\mu, \sigma^2) = \frac{d}{dx}\left(\frac{1}{2\pi\sigma^2} exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} \right)$$

$$= \frac{1}{2\pi\sigma^2}\frac{d}{dx} exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}$$

$$= \frac{1}{2\pi\sigma^2} exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} \left(-\frac{1}{2\sigma^2} \right) (2(x-\mu))$$

Setting it to 0, we get:

$$0 = \frac{1}{2\pi\sigma^2} exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} \left(-\frac{1}{2\sigma^2} \right) (2(x-\mu))$$

$$\implies x = \mu$$

Now we show that the mode of the multi-variate Gaussian distribution is given by $\boldsymbol{\mu}$.

The distribution is given by:

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\mathbf{\Sigma}|^{1/2}} exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\}$$

We take the derivative of this function w.r.t $\mathbf{x}$. This gives us:

$$\frac{d}{d\mathbf{x}}\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \Sigma) = \frac{d}{d\mathbf{x}}\left(\frac{1}{(2\pi)^{D/2}} \frac{1}{|\mathbf{\Sigma}|^{1/2}} exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\} \right)$$

$$= \frac{1}{(2\pi)^{D/2}} \frac{1}{|\mathbf{\Sigma}|^{1/2}} \frac{d}{d\mathbf{x}} \left(exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\} \right)$$

$$= \frac{1}{(2\pi)^{D/2}} \frac{1}{|\mathbf{\Sigma}|^{1/2}} \frac{d}{d\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)} exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}$$

$$*\frac{d}{d\mathbf{x}}\left(-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

$$= \frac{1}{(2\pi)^{D/2}} \frac{1}{|\mathbf{\Sigma}|^{1/2}} exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\} \frac{d}{d\mathbf{x}}\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

Since $\mathbf{\Sigma}^{-1}$ is symmetric, we can apply (85) from the matrix cookbook, giving us:

$$= \frac{1}{(2\pi)^{D/2}} \frac{1}{|\mathbf{\Sigma}|^{1/2}} exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}\left(-\frac{1}{2}\right)\left(2\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

$$= \frac{1}{(2\pi)^{D/2}} \frac{1}{|\mathbf{\Sigma}|^{1/2}} exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}\left(-\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

Setting this to $\mathbf{0}$ vector, we get:

$$\mathbf{0} = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\mathbf{\Sigma}|^{1/2}} exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}\left(-\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

$$\Longrightarrow \mathbf{x} = \boldsymbol{\mu}$$