Let us also consider multidimensional case for this problem: now each x_i is a d-dimensional vector drawn from the multivariate normal distribution with parameters mean $\mu \in \mathbb{R}^d$ and covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$:

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

Similarly, the log likelihood for this model takes the form:

$$\log p(X|\mu,\Sigma) = -\frac{Nd}{2}\log(2\pi) - \frac{N}{2}\log\det\Sigma - \sum_{i=1}^{N}\frac{1}{2}(x_i-\mu)^T\Sigma^{-1}(x_i-\mu) =$$

$$= -\frac{Nd}{2}\log(2\pi) - \frac{N}{2}\log\det\Sigma - \frac{1}{2}\sum_{i=1}^{N}\left(x_i^T\Sigma^{-1}x_i - \mu^T\Sigma^{-1}x_i - x_i^T\Sigma^{-1}\mu + \mu^T\Sigma^{-1}\mu\right) =$$

$\Big[$ use the fact that Σ is symmetric, thus Σ^{-1} is also symmetric which leads to:

$$\mu^T\Sigma^{-1}x_i = \left(\mu^T\Sigma^{-1}x_i\right)^T = x_i^T\left(\Sigma^{-1}\right)^T\mu = x_i^T\Sigma^{-1}\mu \Big]$$

$$= -\frac{Nd}{2}\log(2\pi) - \frac{N}{2}\log\det\Sigma - \frac{1}{2}\sum_{i=1}^{N}\left(x_i^T\Sigma^{-1}x_i - 2x_i^T\Sigma^{-1}\mu + \mu^T\Sigma^{-1}\mu\right)$$

Now to obtain the MLE for μ, we need to compute the derivative of this expression with respect to vector μ and set it to zero. We will use the following vector differentiation rules:

$$\frac{\partial}{\partial y}(a^Ty) = a \quad \text{for } y \in \mathbb{R}^d, a \in \mathbb{R}^d$$

$$\frac{\partial}{\partial y}(y^TAy) = 2Ay \quad \text{for } y \in \mathbb{R}^d \text{ and symmetric matrix } A \in \mathbb{R}^{d\times d}$$

Applying them to the log likelihood expression, we get:

$$\frac{\partial}{\partial\mu}\log p(X|\mu,\Sigma) = -\frac{1}{2}\sum_{i=1}^{N}\left(-2\Sigma^{-1}x_i + 2\Sigma^{-1}\mu\right) = \sum_{i=1}^{N}\Sigma^{-1}\left(x_i-\mu\right) = 0$$

$$\mu_{ML} = \frac{1}{N}\sum_{i=1}^{N}x_i$$