Gaussian Process Regression

Below I outline the step by step process:

1. Compute the covariance matrix, i.e., $K(X, X)$, $K(X, X_*)$, and $K(X_*, K_*)$
2. Compute $\alpha = [K(X, X) + \sigma^2 I]^{-1} \times y$
3. Compute the mean ($\mu = K(X, X_*) \times \alpha$) and sigma ($\sigma = K(X_*, K_*) - [K(X, X) + \sigma^2 I]^{-1} \times K(X, X_*)$)
4. If optimize, then maximize negative log marginal likelihood ($lml = -\frac{1}{2} y^T [K + \sigma^2 I]^{-1} y - \frac{1}{2} log|K + \sigma^2 I| - \frac{n}{2} log 2\pi$)