Problem setup:

- $\gamma_1, \gamma_2, \gamma_3$ are orthonormal unit vectors (3×1 matrices); $\kappa > 0, \beta > 0$ are scalars.
- γ^T is the transpose of γ.
- v_0 is a 3×1 matrix. A is a 3×3 symmetric matrix. v_0 and A are functions of data and can be precomputed.

Parameters to be estimated: $\Theta = \{\kappa, \beta, \gamma_1, \gamma_2, \gamma_3\}$

The negative log-likelihood function that needs to be minimized is as follows:

$$L(\Theta) = \log c(\kappa, \beta) - \kappa \gamma_1^T v_0 - \beta \gamma_2^T A \gamma_2 + \beta \gamma_3^T A \gamma_3 \tag{0.1}$$

where

$$c(\kappa, \beta) = 2\pi \sum_{j=0}^{\infty} \frac{\Gamma(j + \frac{1}{2})}{\Gamma(j+1)} \beta^{2j} \left(\frac{\kappa}{2}\right)^{-2j-\frac{1}{2}} I_{2j+\frac{1}{2}}(\kappa) \tag{0.2}$$

$c(\kappa, \beta)$ is the normalization constant which is an infinite summation and is dependent on the Gamma function (Γ) and the modified Bessel function of the first kind $I_v(\kappa)$.

Equation 0.1 should be minimized subject to the following constraints:

$$\begin{aligned}
\gamma_1^T \gamma_1 &= 1 \\
\gamma_2^T \gamma_2 &= 1 \\
\gamma_3^T \gamma_3 &= 1 \\
\gamma_1^T \gamma_2 &= 0 \\
\gamma_1^T \gamma_3 &= 0
\end{aligned}$$

This problem can be transformed into an unconstrained optimization problem as follows:

- γ_1 is a vector on the unit sphere which can be defined by two angles (co-latitude and longitude parameters) in the spherical coordinate system.
- γ_2 is a unit vector and must be orthogonal to γ_1. Hence, γ_2 is uniquely defined using one parameter.
- γ_3 is then implicitly defined as $\gamma_1 \times \gamma_2$

I am solving this unconstrained problem using the *dlib* library. Explicit gradient vector and Hessian matrix expressions are cumbersome functions. Hence, I am approximating the derivative information by using the utilities provided in the library.