

2.60 Since there are n_i data points belonging to region i , all with probability density h_i , the likelihood function is given by:

$$p(\mathbf{X}|\mathbf{h}) = \prod_{n=1}^N p(\mathbf{x}_n|\mathbf{h}) = \prod_i h_i^{n_i}$$

The log-likelihood is given by:

$$\ln p(\mathbf{X}|\mathbf{h}) = \sum_i n_i \ln h_i$$

We have the normalization constraint:

$$\sum_i h_i \Delta_i = 1$$

And the non-negative constraint on h_i :

$$h_i \geq 0$$

Using Lagrange multipliers to enforce both constraints, we get the following Lagrangian:

$$\mathcal{L}(\dots h_i \dots, \lambda, \dots \gamma_i \dots) = \sum_i n_i \ln h_i + \lambda \left(\sum_i h_i \Delta_i - 1 \right) + \sum_i \gamma_i h_i$$

KKT conditions require:

1. Stationarity

$$\frac{\partial \mathcal{L}(\dots h_i \dots, \lambda, \dots \gamma_i \dots)}{\partial h_i} = 0, \forall i$$

$$\frac{\partial \mathcal{L}(\dots h_i \dots, \lambda, \dots \gamma_i \dots)}{\partial h_i} = \frac{n_i}{h_i} + \lambda \Delta_i + \gamma_i$$

Setting this derivative to 0, we get:

$$\begin{aligned}\frac{n_i}{h_i} + \lambda \Delta_i + \gamma_i &= 0 \\ \implies \lambda h_i \Delta_i + \gamma_i h_i &= -n_i \\ \implies h_i^* &= \frac{-n_i}{\lambda \Delta_i + \gamma_i}\end{aligned}$$

2. Complementary slackness

$$\gamma_i h_i = 0$$

3. Primal feasibility

$$\sum_i h_i^* \Delta_i = 1$$

and

$$h_i^* \geq 0$$

4. Dual feasibility

$$\gamma_i \geq 0$$

Both constraints γ_i and λ cannot be inactive at the same time, because then the denominator in the expression for h_i^* would become 0.

If $\gamma_i > 0$, then due to complementary slackness, $h_i = 0$, regardless of the value of λ .

If $\gamma = 0$ and $\lambda \neq 0$,

$$\lambda h_i^* \Delta_i + \gamma_i h_i = \lambda h_i^* \Delta_i = -n_i$$

$$\implies h_i^* = \frac{-n_i}{\lambda \Delta_i}$$

Applying the primal feasibility condition $\sum_i h_i^* \Delta_i = 1$,

$$\implies \sum_i \frac{-n_i}{\lambda \Delta_i} \Delta_i = 1$$

$$\implies \sum_i \frac{-n_i}{\lambda} = 1$$

$$\implies \lambda = - \sum_i n_i$$

$$\implies \lambda = -N$$

Substituting this value of λ back into the solution for h_i^* ,

$$h_i^* = \frac{n_i}{N \Delta_i}$$

To check that the solution is a global optimum, the following four conditions must hold (Source):

- Objective function should be concave.
 $\sum_i n_i \ln h_i$ is a sum of log functions which are all concave, so the objective function is also concave.
- The inequality constraints should be differentiable convex functions.
The inequality constraints are $-h_i \forall i$, which are both differentiable and convex.
- The equality constraints should be affine functions.
The equality constraint is $\sum_i h_i \Delta_i - 1$, which is a linear function with a constant and is therefore affine.
- Slater's condition should hold.
 $h_i^* = \frac{n_i}{N \Delta_i} \forall i$, satisfies all constraints and there are no non-linear constraints. So, Slater's condition holds. Slater's Condition.

Since these conditions are satisfied, the KKT conditions guarantee a global optimum.