Group 6

Overview

The proposed approach consists in minimizing the number of ASVs that would fail to fall into the normal abundance range after the treatment. It is assumed that after the intervention each ASV’s abundance is increased by **b** = **NIM** × **x**, where **x** encodes both the set of treatment nutrients and their impact (more precisely, increase) on the ASVs’ abundances.

Questions

1. **NIM** is a matrix of 0s and 1s. Vector **x** has only positive values. Thus, vector **b** = **NIM** × **x** would have only positive values. When computing **c**, the original abundance of ASVs (i.e., MS) is being increased **b**. This abundance was normalized to sum up to 1 in each sample. However, after adding **b** to **MS**, the total abundance in each sample would be greater than 1, i.e., not normalized. Does this have any implications?
2. In the formula for **c**, the chosen measure of impact is additive (namely the addition of **b**, a value that does not depend on the initial ASV abundance). But what if the underlying biology implies scaling of the original abundance by some factor, e.g., multiplication by **b** or **1+b**? What implications might it have? (For example, in the additive approach, ASV abundances 0.001 in sample A and 0.00001 in sample B would grow by the same amount, let’s say, b=0.05. Is that OK?)
3. Normal range is being defined by some thresholds. What implications these thresholds may have? What may the equivalent continuous metric look like, i.e., the one without categorizing abundances into normal/abnormal using thresholds?